

On cyclic DNA codes over the rings $Z_4 + wZ_4$ and $Z_4 + wZ_4 + vZ_4 + wvZ_4$

Abdullah Dertli^a, Yasemin Cengellenmis^b

(a) Ondokuz Mayıs University, Faculty of Arts and Sciences,
Mathematics Department, Samsun, Turkey
abdullah.dertli@gmail.com

(b) Trakya University, Faculty of Arts and Sciences,
Mathematics Department, Edirne, Turkey
ycengellenmis@gmail.com

May 11, 2016

Abstract

The structures of cyclic DNA codes of odd length over the finite rings $R = Z_4 + wZ_4$, $w^2 = 2$ and $S = Z_4 + wZ_4 + vZ_4 + wvZ_4$, $w^2 = 2$, $v^2 = v$, $wv = vw$ are studied. The links between the elements of the rings R , S and 16 and 256 codons are established, respectively. Cyclic codes of odd length over the finite ring R satisfies reverse complement constraint and cyclic codes of odd length over the finite ring S satisfy reverse constraint and reverse complement constraint are studied. Binary images of the cyclic DNA codes over the finite rings R and S are determined. Moreover, a family of DNA skew cyclic codes over R is constructed, its property of being reverse complement is studied.

1 Introduction

DNA is formed by the strands and each strands is sequence consists of four nucleotides ; adenine (A), guanine (G), thymine (T) and cytosine (C). Two strands of DNA are linked with Watson-Crick Complement. This is as $\overline{A} = T$, $\overline{T} = A$, $\overline{G} = C$, $\overline{C} = G$. For example if $c = (ATCCG)$ then its complement is $\overline{c} = (TAGGC)$.

A code is called DNA codes if it satisfies some or all of the following conditions;

- i) The Hamming constraint, for any two different codewords $c_1, c_2 \in C$,
 $H(c_1, c_2) \geq d$
- ii) The reverse constraint, for any two different codewords $c_1, c_2 \in C$,
 $H(c_1, c_2^r) \geq d$

iii) The reverse complement constraint, for any two different codewords $c_1, c_2 \in C$, $H(c_1, c_2^c) \geq d$

iv) The fixed GC content constraint, for any codeword $c \in C$ contains the same number of G and C element.

The purpose of the i-iii constraints is to avoid undesirable hybridization between different strands.

DNA computing were started by Leonhard Adleman in 1994, in [3]. The special error correcting codes over some finite fields and finite rings with 4^n elements where $n \in \mathbb{N}$ were used for DNA computing applications.

In [12], the reversible codes over finite fields were studied, firstly. It was shown that $C = \langle f(x) \rangle$ is reversible if and only if $f(x)$ is a self reciprocal polynomial. In [1], they developed the theory for constructing linear and additive cyclic codes of odd length over $GF(4)$. In [13], they introduced a new family of polynomials which generates reversible codes over a finite field $GF(16)$.

In [2], the reversible cyclic codes of any length n over the ring Z_4 were studied. A set of generators for cyclic codes over Z_4 with no restrictions on the length n was found. In [17], cyclic DNA codes over the ring $R = \{0, 1, u, 1+u\}$ where $u^2 = 1$ based on a similarity measure were constructed. In [9], the codes over the ring $F_2 + uF_2, u^2 = 0$ were constructed for using in DNA computing applications.

I. Siap et al. considered cyclic DNA codes over the finite ring $F_2[u]/\langle u^2 - 1 \rangle$ in [18]. In [10], Liang and Wang considered cyclic DNA codes over $F_2 + uF_2, u^2 = 0$. Yıldız and Siap studied cyclic DNA codes over $F_2[u]/\langle u^4 - 1 \rangle$ in [19]. Bayram et al. considered codes over the finite ring $F_4 + vF_4, v^2 = v$ in [3]. Zhu and Chan studied cyclic DNA codes over the non-chain ring $F_2[u, v]/\langle u^2, v^2 - v, uv - vu \rangle$ in [20]. In [5], Bemenni et al. studied cyclic DNA codes over $F_2[u]/\langle u^6 \rangle$. Pattanayak et al. considered cyclic DNA codes over the ring $F_2[u, v]/\langle u^2 - 1, v^3 - v, uv - vu \rangle$ in [15]. Pattanayak and Singh studied cyclic DNA codes over the ring $Z_4 + uZ_4, u^2 = 0$ in [14].

J. Gao et al. studied the construction of cyclic DNA codes by cyclic codes over the finite ring $F_4[u]/\langle u^2 + 1 \rangle$, in [11]. Also, the construction of DNA cyclic codes have been discussed by several authors in [7, 8, 16].

We study families of DNA cyclic codes of the finite rings $Z_4 + wZ_4, w^2 = 2$ and $Z_4 + wZ_4 + vZ_4 + wvZ_4, w^2 = 2, v^2 = v, wv = vw$. The rest of the paper is organized as follows. In section 2, details about algebraic structure of the finite ring $Z_4 + wZ_4, w^2 = 2$ are given. We define a Gray map from R to Z_4 . In section 3, cyclic codes of odd length over R satisfies the reverse complement constraint are determined. In section 4, cyclic codes of odd length over S satisfy the reverse complement constraint and the reverse constraint are examined. A linear code over S is represented by means of two linear codes over R . In section 5, the binary image of cyclic DNA code over R is determined. In section 6, the binary image of cyclic DNA code over S is determined. In section 7, by using a non trivial automorphism, the DNA skew cyclic codes are introduced. In section 8, the design of linear DNA code is presented.

2 Preliminaries

The algebraic structure of the finite ring $R = Z_4 + wZ_4$, $w^2 = 2$ is given in [6]. R is the commutative, characteristic 4 ring $Z_4 + wZ_4 = \{a + wb : a, b \in Z_4\}$ with $w^2 = 2$. R can also be thought of as the quotient ring $Z_4[w]/\langle w^2 - 2 \rangle$. R is principal ideal ring with 16 elements and finite chain ring. The units of the ring are

$$1, 3, 1 + w, 3 + w, 1 + 2w, 1 + 3w, 3 + 3w, 3 + 2w$$

and the non units are

$$0, 2, w, 2w, 3w, 2 + w, 2 + 2w, 2 + 3w$$

R has 4 ideals in all

$$\begin{aligned} \langle 0 \rangle &= \{0\} \\ \langle 1 \rangle &= \langle 3 \rangle = \langle 1 + 3w \rangle = \dots = R \\ \langle w \rangle &= \{0, 2, w, 2w, 3w, 2 + w, 2 + 2w, 2 + 3w\} \\ &= \langle 3w \rangle = \langle 2 + w \rangle = \langle 2 + 3w \rangle \\ \langle 2w \rangle &= \{0, w\} \\ \langle 2 \rangle &= \langle 2 + 2w \rangle = \{0, 2, 2w, 2 + 2w\} \end{aligned}$$

$$\langle 0 \rangle \subset \langle 2w \rangle \subset \langle 2 \rangle \subset \langle w \rangle \subset R$$

Moreover R is Frobenious ring.

We define

$$\begin{aligned} \phi &: R \longrightarrow Z_4^2 \\ \phi(a + wb) &= (a, b) \end{aligned}$$

This Gray map is extended component wise to

$$\begin{aligned} \phi &: R^n \longrightarrow Z_4^{2n} \\ (\alpha_1, \alpha_2, \dots, \alpha_n) &= (a_1, \dots, a_n, b_1, \dots, b_n) \end{aligned}$$

where $\alpha_i = a_i + b_i w$ with $i = 1, 2, \dots, n$. ϕ is Z_4 module isomorphism.

A linear code C of length n over R is a R -submodule of R^n . An element of C is called a codeword. A code of length n is cyclic if the code invariant under the automorphism σ which

$$\sigma(c_0, c_1, \dots, c_{n-1}) = (c_{n-1}, c_0, \dots, c_{n-2})$$

A cyclic code of length n over R can be identified with an ideal in the quotient ring $R[x]/\langle x^n - 1 \rangle$ via the R -modul isomorphism

$$\begin{aligned} R^n &\longrightarrow R[x]/\langle x^n - 1 \rangle \\ (c_0, c_1, \dots, c_{n-1}) &\longmapsto c_0 + c_1 x + \dots + c_{n-1} x^{n-1} + \langle x^n - 1 \rangle \end{aligned}$$

Theorem 1 Let C be a cyclic code in $R[x]/\langle x^n - 1 \rangle$. Then there exists polynomials $g(x), a(x)$ such that $a(x)|g(x)|x^n - 1$ and $C = \langle g(x), wa(x) \rangle$.

The ring $R[x]/\langle x^n - 1 \rangle$ is a principal ideal ring when n is odd. So, if n is odd, then there exist $s(x) \in R[x]/\langle x^n - 1 \rangle$ such that $C = \langle s(x) \rangle$.

3 The reversible complement codes over R

In this section, we study cyclic code of odd length over R satisfies the reverse complement constraint. Let $\{A, T, G, C\}$ represent the DNA alphabet. DNA occurs in sequences with represented by sequences of the DNA alphabet. DNA code of length n is defined as a set of the codewords $(x_0, x_1, \dots, x_{n-1})$ where $x_i \in \{A, T, G, C\}$. These codewords must satisfy the four constraints which are mentioned in [20].

Since the ring R is of cardinality 16, we define the map ϕ which gives a one to one correspondence between the elements of R and the 16 codons over the alphabet $\{A, T, G, C\}^2$ by using the Gray map as follows

Elements	Gray images	DNA double pairs
0	(0, 0)	AA
1	(1, 0)	CA
2	(2, 0)	GA
3	(3, 0)	TA
w	(0, 1)	AC
$2w$	(0, 2)	AG
$3w$	(0, 3)	AT
$1 + w$	(1, 1)	CC
$1 + 2w$	(1, 2)	CG
$1 + 3w$	(1, 3)	CT
$2 + w$	(2, 1)	GC
$2 + 2w$	(2, 2)	GG
$2 + 3w$	(2, 3)	GT
$3 + w$	(3, 1)	TC
$3 + 2w$	(3, 2)	TG
$3 + 3w$	(3, 3)	TT

The codons satisfy the Watson-Crick Complement.

Definition 2 For $x = (x_0, x_1, \dots, x_{n-1}) \in R^n$, the vector $(x_{n-1}, x_{n-2}, \dots, x_1, x_0)$ is called the reverse of x and is denoted by x^r . A linear code C of length n over R is said to be reversible if $x^r \in C$ for every $x \in C$.

For $x = (x_0, x_1, \dots, x_{n-1}) \in R^n$, the vector $(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_{n-1})$ is called the complement of x and is denoted by x^c . A linear code C of length n over R is said to be complement if $x^c \in C$ for every $x \in C$.

For $x = (x_0, x_1, \dots, x_{n-1}) \in R^n$, the vector $(\bar{x}_{n-1}, \bar{x}_{n-2}, \dots, \bar{x}_1, \bar{x}_0)$ is called the reversible complement of x and is denoted by x^{rc} . A linear code C of length n over R is said to be reversible complement if $x^{rc} \in C$ for every $x \in C$.

Definition 3 Let $f(x) = a_0 + a_1x + \dots + a_rx^r$ with $a_r \neq 0$ be polynomial. The reciprocal of $f(x)$ is defined as $f^*(x) = x^r f(\frac{1}{x})$. It is easy to see that $\deg f^*(x) \leq \deg f(x)$ and if $a_0 \neq 0$, then $\deg f^*(x) = \deg f(x)$. $f(x)$ is called a self reciprocal polynomial if there is a constant m such that $f^*(x) = mf(x)$.

Lemma 4 Let $f(x), g(x)$ be polynomials in $R[x]$. Suppose $\deg f(x) - \deg g(x) = m$ then,

- i) $(f(x)g(x))^* = f^*(x)g^*(x)$
- ii) $(f(x) + g(x))^* = f^*(x) + x^m g^*(x)$

Lemma 5 For any $a \in R$, we have $a + \bar{a} = 3 + 3w$.

Lemma 6 If $a \in \{0, 1, 2, 3\}$, then we have $(3 + 3w) - \bar{a} = wa$.

Theorem 7 Let $C = \langle g(x), wa(x) \rangle$ be a cyclic code of odd length n over R . If $f(x)^{rc} \in C$ for any $f(x) \in C$, then $(1 + w)(1 + x + x^2 + \dots + x^{n-1}) \in C$ and there are two constants $e, d \in Z_4^*$ such that $g^*(x) = eg(x)$ and $a^*(x) = da(x)$.

Proof. Suppose that $C = \langle g(x), wa(x) \rangle$, where $a(x)|g(x)|x^n - 1 \in Z_4[x]$. Since $(0, 0, \dots, 0) \in C$, then its reversible complement is also in C .

$$\begin{aligned} (0, 0, \dots, 0)^{rc} &= (3 + 3w, 3 + 3w, \dots, 3 + 3w) \\ &= 3(1 + w)(1, 1, \dots, 1) \in C \end{aligned}$$

This vector corresponds of the polynomial

$$(3 + 3w) + (3 + 3w)x + \dots + (3 + 3w)x^{n-1} = (3 + 3w)\frac{x^n - 1}{x - 1} \in C$$

Since $3 \in Z_4^*$, then $(1 + w)(1 + x + \dots + x^{n-1}) \in C$.

Let $g(x) = g_0 + g_1x + \dots + g_{r-1}x^{r-1} + g_rx^r$. Note that $g(x)^{rc} = (3 + 3w) + (3 + 3w)x + \dots + (3 + 3w)x^{n-r-2} + \bar{g}_rx^{n-r-1} + \dots + \bar{g}_1x^{n-2} + \bar{g}_0x^{n-1} \in C$.

Since C is a linear code, then

$$3(1 + w)(1 + x + x^2 + \dots + x^{n-1}) - g(x)^{rc} \in C$$

which implies that $((3 + 3w) - \bar{g}_r)x^{n-r-1} + ((3 + 3w) - \bar{g}_{r-1})x^{n-r-2} + \dots + ((3 + 3w) - \bar{g}_0)x^{n-1} \in C$. By using $(3 + 3w) - \bar{a} = a$, this implies that

$$x^{n-r-1}(g_r + g_{r-1}x + \dots + g_0x^r) = x^{n-r-1}g^*(x) \in C$$

Since $g^*(x) \in C$, this implies that

$$g^*(x) = g(x)u(x) + wa(x)v(x)$$

where $u(x), v(x) \in Z_4[x]$. Since $g_i \in Z_4$ for $i = 0, 1, \dots, r$, we have that $v(x) = 0$. As $\deg g^*(x) = \deg g(x)$, we have $u(x) \in Z_4^*$. Therefore there is a constant $e \in Z_4^*$ such that $g^*(x) = eg(x)$. So, $g(x)$ is a self reciprocal polynomial.

Let $a(x) = a_0 + a_1x + \dots + a_tx^t$. Suppose that $wa(x) = wa_0 + wa_1x + \dots + wa_tx^t$. Then

$$(wa(x))^{rc} = (3+3w) + (3+3w)x + \dots + \overline{wa_t}x^{n-t-1} + \dots + \overline{wa_1}x^{n-2} + \overline{wa_0}x^{n-1} \in C$$

As $(3+3w)\frac{x^n-1}{x-1} \in C$ and C is a linear code, then

$$-(wa(x))^{rc} + (3+3w)\frac{x^n-1}{x-1} \in C$$

Hence, $x^{n-t-1}[(-\overline{wa_t}) + (3+3w)] + (-\overline{wa_{t-1}}) + (3+3w)x + \dots + (-\overline{wa_0}) + (3+3w)x^t]$. By Lemma 6, we get

$$x^{n-t-1}(wa_t + wa_{t-1}x + \dots + wa_0x^t)$$

$x^{n-t-1}wa^*(x) \in C$. Since $wa^*(x) \in C$, we have

$$wa^*(x) = g(x)h(x) + wa(x)s(x)$$

Since w doesn't appear in $g(x)$, it follows that $h(x) = 0$ and $a^*(x) = a(x)s(x)$. As $\deg a^*(x) = \deg a(x)$, then $s(x) \in Z_4^*$. So, $a(x)$ is a self reciprocal polynomial. ■

Theorem 8 Let $C = \langle g(x), wa(x) \rangle$ be a cyclic code of odd length n over R . If $(1+w)(1+x+x^2+\dots+x^{n-1}) \in C$ and $g(x), a(x)$ are self reciprocal polynomials, then $c(x)^{rc} \in C$ for any $c(x) \in C$.

Proof. Since $C = \langle g(x), wa(x) \rangle$, for any $c(x) \in C$, there exist $m(x)$ and $n(x)$ in $R[x]$ such that $c(x) = g(x)m(x) + wa(x)n(x)$. By using Lemma 4, we have

$$\begin{aligned} c^*(x) &= (g(x)m(x) + wa(x)n(x)) \\ &= (g(x)m(x))^* + x^s(wa(x)n(x)) \\ &= g^*(x)m^*(x) + wa^*(x)(x^sn^*(x)) \end{aligned}$$

Since $g^*(x) = eg(x)$, $a^*(x) = da(x)$, we have $c^*(x) = eg(x)m^*(x) + dwa(x)(x^sn^*(x)) \in C$. So, $c^*(x) \in C$.

Let $c(x) = c_0 + c_1x + \dots + c_tx^t \in C$. Since C is a cyclic code, we get

$$x^{n-t-1}c(x) = c_0x^{n-t-1} + c_1x^{n-t} + \dots + c_tx^{n-1} \in C$$

Since $(1+w) + (1+w)x + \dots + (1+w)x^{n-1} \in C$ and C is a linear code

$$\begin{aligned} -(1+w)\frac{x^n-1}{x-1} - x^{n-t-1}c(x) &= -(1+w) - (1+w)x + \dots + (-c_0 - (1+w))x^{n-t-1} \\ &\quad + \dots + (-c_t - (1+w))x^{n-1} \in C \end{aligned}$$

By using $\bar{a} + (1+w) = -a$, this implies that

$$-(1+w) - \dots + \bar{c}_0x^{n-t-1} + \dots + \bar{c}_tx^{n-1} \in C$$

This shows that $(c^*(x))^{rc} \in C$.

$$((c^*(x))^{rc})^* = \bar{c}_t + \bar{c}_{t-1}x + \dots + (3+3w)x^{n-1}$$

This corresponds this vector $(\bar{c}_t, \bar{c}_{t-1}, \dots, \bar{c}_0, \dots, \bar{0})$. Since $(c^*(x)^{rc})^* = (x^{n-t-1}c(x))^{rc}$, so $c(x)^{rc} \in C$. ■

4 The reversible and reversible complement codes over S

Throughout this paper, S denotes the commutative ring $Z_4 + wZ_4 + vZ_4 + wvZ_4 = \{b_1 + wb_2 + vb_3 + wvb_4 : b_j \in Z_4, 1 \leq j \leq 4\}$ with $w^2 = 2, v^2 = v, wv = vw$, with characteristic 4. S can also be thought of as the quotient ring $Z_4[w, v] / \langle w^2 - 2, v^2 - v, wv - vw \rangle$.

Let

$$\begin{aligned} S &= Z_4 + wZ_4 + vZ_4 + wvZ_4, \text{ where } w^2 = 2, v^2 = v, wv = vw \\ &= (Z_4 + wZ_4) + v(Z_4 + wZ_4), \text{ where } w^2 = 2, v^2 = v, wv = vw \\ &= R + vR, \text{ where } v^2 = v \end{aligned}$$

We define the Gray map ϕ_1 from S to R as follows

$$\begin{aligned} \phi_1 &: S \longrightarrow R^2 \\ a + vb &\longmapsto (a, b) \end{aligned}$$

where $a, b \in R$. This Gray map is extended componentwise to

$$\begin{aligned} \phi_1 &: S^n \longrightarrow R^{2n} \\ x &= (x_1, \dots, x_n) \longmapsto (a_1, \dots, a_n, b_1, \dots, b_n) \end{aligned}$$

where $x_i = a_i + vb_i, a_i, b_i \in R$ for $i = 1, 2, \dots, n$.

In this section, we study cyclic codes of odd length n over S satisfy reverse and reverse complement constraint. Since the ring S is of the cardinality 4^4 , then we define the map ϕ_1 which gives a one to one correspondence between the element of S and the 256 codons over the alphabet $\{A, T, G, C\}^4$ by using the Gray map. For example;

$$\begin{aligned} 0 &= 0 + v0 \longmapsto \phi_1(0) = (0, 0) \longrightarrow AAAA \\ 2wv &= 0 + v(2w) \longmapsto \phi_1(2wv) = (0, 2w) \longrightarrow AAAG \\ 1 + 3v + 3wv &= 1 + v(3 + 3w) \longmapsto \phi_1(1 + v(3 + 3w)) = (1, 3 + 3w) \longrightarrow CATT \end{aligned}$$

Definition 9 Let A_1, A_2 be linear codes.

$$A_1 \otimes A_2 = \{(a_1, a_2) : a_1 \in A_1, a_2 \in A_2\}$$

and

$$A_1 \oplus A_2 = \{a_1 + a_2 : a_1 \in A_1, a_2 \in A_2\}$$

Let C be a linear code of length n over S . Define

$$\begin{aligned} C_1 &= \{a : \exists b \in R^n, a + vb \in C\} \\ C_2 &= \{b : \exists a \in R^n, a + vb \in C\} \end{aligned}$$

where C_1 and C_2 are linear codes over R of length n .

Theorem 10 Let C be a linear code of length n over S . Then $\phi_1(C) = C_1 \otimes C_2$ and $|C| = |C_1| |C_2|$.

Corollary 11 If $\phi_1(C) = C_1 \otimes C_2$, then $C = vC_1 \oplus (1-v)C_2$.

Theorem 12 Let $C = vC_1 \oplus (1-v)C_2$ be a linear code of odd length n over S . Then C is a cyclic code over S if and only if C_1, C_2 are cyclic codes over R .

Proof. Let $(a_0^1, a_1^1, \dots, a_{n-1}^1) \in C_1, (a_0^2, a_1^2, \dots, a_{n-1}^2) \in C_2$. Assume that $m_i = va_i^1 \oplus (1-v)a_i^2$ for $i = 0, 1, 2, \dots, n-1$. Then $(m_0, m_1, \dots, m_{n-1}) \in C$. Since C is a cyclic code, it follows that $(m_{n-1}, m_0, m_1, \dots, m_{n-2}) \in C$. Note that $(m_{n-1}, m_0, \dots, m_{n-2}) = v(a_{n-1}^1, a_0^1, \dots, a_{n-2}^1) \oplus (1-v)(a_{n-1}^2, a_0^2, \dots, a_{n-2}^2)$. Hence $(a_{n-1}^1, a_0^1, \dots, a_{n-2}^1) \in C_1, (a_{n-1}^2, a_0^2, \dots, a_{n-2}^2) \in C_2$. Therefore C_1, C_2 are cyclic codes over R .

Conversely, suppose that C_1, C_2 are cyclic codes over R . Let $(m_0, m_1, \dots, m_{n-1}) \in C$, where $m_i = va_i^1 \oplus (1-v)a_i^2$ for $i = 0, 1, 2, \dots, n-1$. Then $(a_{n-1}^1, a_0^1, \dots, a_{n-2}^1) \in C_1, (a_{n-1}^2, a_0^2, \dots, a_{n-2}^2) \in C_2$. Note that $(m_{n-1}, m_0, \dots, m_{n-2}) = v(a_{n-1}^1, a_0^1, \dots, a_{n-2}^1) \oplus (1-v)(a_{n-1}^2, a_0^2, \dots, a_{n-2}^2) \in C$. So, C is a cyclic code over S . ■

Theorem 13 Let $C = vC_1 \oplus (1-v)C_2$ be a linear code of odd length n over S . Then C is reversible over S iff C_1, C_2 are reversible over R .

Proof. Let C_1, C_2 be reversible codes. For any $b \in C, b = vb_1 + (1-v)b_2$, where $b_1 \in C_1, b_2 \in C_2$. Since C_1 and C_2 are reversible, $b_1^r \in C_1, b_2^r \in C_2$. So, $b^r = vb_1^r + (1-v)b_2^r \in C$. Hence C is reversible.

On the other hand, Let C be a reversible code over S . So for any $b = vb_1 + (1-v)b_2 \in C$, where $b_1 \in C_1, b_2 \in C_2$, we get $b^r = vb_1^r + (1-v)b_2^r \in C$. Let $b^r = vb_1^r + (1-v)b_2^r = vs_1 + (1-v)s_2$, where $s_1 \in C_1, s_2 \in C_2$. So C_1 and C_2 are reversible codes over R . ■

Lemma 14 For any $c \in S$, we have $c + \bar{c} = (3 + 3w) + v(3 + 3w)$.

Lemma 15 For any $a \in S, \bar{a} + 3\bar{0} = 3a$.

Theorem 16 Let $C = vC_1 \oplus (1-v)C_2$ be a cyclic code of odd length n over S . Then C is reversible complement over S iff C is reversible over S and $(\bar{0}, \bar{0}, \dots, \bar{0}) \in C$.

Proof. Since C is reversible complement, for any $c = (c_0, c_1, \dots, c_{n-1}) \in C, c^{rc} = (\bar{c}_{n-1}, \bar{c}_{n-2}, \dots, \bar{c}_0) \in C$. Since C is a linear code, so $(0, 0, \dots, 0) \in C$. Since C is reversible complement, so $(\bar{0}, \bar{0}, \dots, \bar{0}) \in C$. By using Lemma 15, we have

$$3c^r = 3(c_{n-1}, c_{n-2}, \dots, c_0) = (\bar{c}_{n-1}, \bar{c}_{n-2}, \dots, \bar{c}_0) + 3(\bar{0}, \bar{0}, \dots, \bar{0}) \in C$$

So, for any $c \in C$, we have $c^r \in C$.

On the other hand, let C be reversible. So, for any $c = (c_0, c_1, \dots, c_{n-1}) \in C, c^r = (c_{n-1}, c_{n-2}, \dots, c_0) \in C$. To show that C is reversible complement, for any $c \in C$,

$$c^{rc} = (\bar{c}_{n-1}, \bar{c}_{n-2}, \dots, \bar{c}_0) = 3(c_{n-1}, c_{n-2}, \dots, c_0) + (\bar{0}, \bar{0}, \dots, \bar{0}) \in C$$

So, C is reversible complement. ■

Lemma 17 For any $a, b \in S$, $\overline{a+b} = \overline{a} + \overline{b} - 3(1+w)(1+v)$.

Theorem 18 Let D_1 and D_2 be two reversible complement cyclic codes of length n over S . Then $D_1 + D_2$ and $D_1 \cap D_2$ are reversible complement cyclic codes.

Proof. Let $d_1 = (c_0, c_1, \dots, c_{n-1}) \in D_1, d_2 = (c_0^1, c_1^1, \dots, c_{n-1}^1) \in D_2$. Then,

$$\begin{aligned}
(d_1 + d_2)^{rc} &= \left(\overline{(c_{n-1} + c_{n-1}^1)}, \dots, \overline{(c_1 + c_1^1)}, \overline{(c_0 + c_0^1)} \right) \\
&= \left(\overline{c_{n-1}} + \overline{c_{n-1}^1} - 3(1+w)(1+v), \dots, \overline{c_0} + \overline{c_0^1} - 3(1+w)(1+v) \right) \\
&= \left(\overline{c_{n-1}} - 3(1+w)(1+v), \dots, \overline{c_0} - 3(1+w)(1+v) \right) + \left(\overline{c_{n-1}^1}, \dots, \overline{c_0^1} \right) \\
&= \left(d_1^{rc} - 3(1+w)(1+v) \frac{x^n - 1}{x - 1} \right) + d_2^{rc} \in D_1 + D_2
\end{aligned}$$

This shows that $D_1 + D_2$ is reversible complement cyclic code. It is clear that $D_1 \cap D_2$ is reversible complement cyclic code. ■

5 Binary images of cyclic DNA codes over R

The 2-adic expansion of $c \in Z_4$ is $c = \alpha(c) + 2\beta(c)$ such that $\alpha(c) + \beta(c) + \gamma(c) = 0$ for all $c \in Z_4$

c	$\alpha(c)$	$\beta(c)$	$\gamma(c)$
0	0	0	0
1	1	0	1
2	0	1	1
3	1	1	0

The Gray map is given by

$$\begin{aligned}
\Psi &: Z_4 \longrightarrow Z_2^2 \\
c &\longmapsto \Psi(c) = (\beta(c), \gamma(c))
\end{aligned}$$

for all $c \in Z_4$ in [14]. Define

$$\begin{aligned}
\check{O} &: R \longrightarrow Z_4^2 \\
a + bw &\longmapsto \check{O}(a + wb) = \Psi(\phi(a + wb)) \\
&= \Psi(a, b) \\
&= (\beta(a), \gamma(a), \beta(b), \gamma(b))
\end{aligned}$$

Let $a + wb$ be any element of the ring R . The Lee weight w_L of the ring R is defined as follows

$$w_L(a + wb) = w_L(a, b)$$

where $w_L(a, b)$ described the usual Lee weight on Z_4^2 . For any $c_1, c_2 \in R$ the Lee distance d_L is given by $d_L(c_1, c_2) = w_L(c_1 - c_2)$.

The Hamming distance $d(c_1, c_2)$ between two codewords c_1 and c_2 is the Hamming weight of the codewords $c_1 - c_2$.

AA	\longrightarrow	0000	CG	\longrightarrow	0111
CA	\longrightarrow	0100	CT	\longrightarrow	0110
GA	\longrightarrow	1100	GC	\longrightarrow	1101
TA	\longrightarrow	1000	GG	\longrightarrow	1111
AC	\longrightarrow	0001	GT	\longrightarrow	1110
AG	\longrightarrow	0011	TC	\longrightarrow	1001
AT	\longrightarrow	0010	TG	\longrightarrow	1011
CC	\longrightarrow	0101	TT	\longrightarrow	1010

Lemma 19 *The Gray map \check{O} is a distance preserving map from $(R^n, \text{Lee distance})$ to $(Z_2^{4n}, \text{Hamming distance})$. It is also Z_2 -linear.*

Proof. For $c_1, c_2 \in R^n$, we have $\check{O}(c_1 - c_2) = \check{O}(c_1) - \check{O}(c_2)$. So, $d_L(c_1, c_2) = w_L(c_1 - c_2) = w_H(\check{O}(c_1 - c_2)) = w_H(\check{O}(c_1) - \check{O}(c_2)) = d_H(\check{O}(c_1), \check{O}(c_2))$. So, the Gray map \check{O} is distance preserving map. For any $c_1, c_2 \in R^n, k_1, k_2 \in Z_2$, we have $\check{O}(k_1 c_1 + k_2 c_2) = k_1 \check{O}(c_1) + k_2 \check{O}(c_2)$. Thus, \check{O} is Z_2 -linear. ■

Proposition 20 *Let σ be the cyclic shift of R^n and v be the 4-quasi-cyclic shift of Z_2^{4n} . Let \check{O} be the Gray map from R^n to Z_2^{4n} . Then $\check{O}\sigma = v\check{O}$.*

Proof. Let $c = (c_0, c_1, \dots, c_{n-1}) \in R^n$, we have $c_i = a_{1i} + w b_{2i}$ with $a_{1i}, b_{2i} \in Z_4, 0 \leq i \leq n-1$. By applying the Gray map, we have

$$\check{O}(c) = \begin{pmatrix} \beta(a_{10}), \gamma(a_{10}), \beta(b_{20}), \gamma(b_{20}), \beta(a_{11}), \gamma(a_{11}), \beta(b_{21}), \gamma(b_{21}), \dots, \\ \beta(a_{1n-1}), \gamma(a_{1n-1}), \beta(b_{2n-1}), \gamma(b_{2n-1}) \end{pmatrix}.$$

Hence

$$v(\check{O}(c)) = \begin{pmatrix} \beta(a_{1n-1}), \gamma(a_{1n-1}), \beta(b_{2n-1}), \gamma(b_{2n-1}), \beta(a_{10}), \gamma(a_{10}), \beta(b_{20}), \\ \gamma(b_{20}), \dots, \beta(a_{1n-2}), \gamma(a_{1n-2}), \beta(b_{2n-2}), \gamma(b_{2n-2}) \end{pmatrix}.$$

On the other hand, $\sigma(c) = (c_{n-1}, c_0, c_1, \dots, c_{n-2})$. We have

$$\check{O}(\sigma(c)) = \begin{pmatrix} \beta(a_{1n-1}), \gamma(a_{1n-1}), \beta(b_{2n-1}), \gamma(b_{2n-1}), \beta(a_{10}), \gamma(a_{10}), \\ \beta(b_{20}), \gamma(b_{20}), \dots, \beta(a_{1n-2}), \gamma(a_{1n-2}), \beta(b_{2n-2}), \gamma(b_{2n-2}) \end{pmatrix}.$$

Therefore, $\check{O}\sigma = v\check{O}$. ■

Theorem 21 *If C is a cyclic DNA code of length n over R then \check{O} is a binary quasi-cyclic DNA code of length $4n$ with index 4.*

6 Binary image of cyclic DNA codes over S

We define

$$\begin{aligned} \tilde{\Psi} &: S \longrightarrow Z_4^4 \\ a_0 + w a_1 + v a_2 + w v a_3 &\longmapsto (a_0, a_1, a_2, a_3) \end{aligned}$$

where $a_i \in Z_4$ for $i = 0, 1, 2, 3$.

Now, we define

$$\begin{aligned} \Theta &: S \longrightarrow Z_2^8 \\ a_0 + wa_1 + va_2 + wva_3 &\longmapsto \Theta(a_0 + wa_1 + va_2 + wva_3) = \Psi(\tilde{\Psi}(a_0 + wa_1 + va_2 + wva_3)) \\ &= (\beta(a_0), \gamma(a_0), \beta(a_1), \gamma(a_1), \beta(a_2), \gamma(a_2), \beta(a_3), \gamma(a_3)) \end{aligned}$$

where Ψ is the Gray map Z_4 to Z_2^2 .

Let $a_0 + wa_1 + va_2 + wva_3$ be any element of the ring S . The Lee weight w_L of the ring S is defined as

$$w_L(a_0 + wa_1 + va_2 + wva_3) = w_L((a_0, a_1, a_2, a_3))$$

where $w_L((a_0, a_1, a_2, a_3))$ described the usual Lee weight on Z_4^4 . For any $c_1, c_2 \in S$, the Lee distance d_L is given by $d_L(c_1, c_2) = w_L(c_1 - c_2)$.

The Hamming distance $d(c_1, c_2)$ between two codewords c_1 and c_2 is the Hamming weight of the codewords $c_1 - c_2$.

Binary image of the codons;

$$\begin{array}{lll} AAAA & \longrightarrow & 00000000 \\ AAC A & \longrightarrow & 00000100 \\ AAG A & \longrightarrow & 00001100 \\ AAT A & \longrightarrow & 00001000 \\ \vdots & \vdots & \vdots \end{array}$$

Lemma 22 *The Gray map Θ is a distance preserving map from $(S^n, \text{Lee distance})$ to $(Z_2^{8n}, \text{Hamming distance})$. It is also Z_2 -linear.*

Proof. It is proved as in the proof of Lemma 19. ■

Proposition 23 *Let σ be the cyclic shift of S^n and $\overset{\circ}{v}$ be the 8-quasi-cyclic shift of Z_2^{8n} . Let Θ be the Gray map from S^n to Z_2^{8n} . Then $\Theta\sigma = \overset{\circ}{v}\Theta$.*

Proof. It is proved as in the proof of Proposition 20. ■

Theorem 24 *If C is a cyclic DNA code of length n over S then Θ is a binary quasi-cyclic DNA code of length $8n$ with index 8.*

Proof. Let C be a cyclic DNA code of length n over S . So, $\sigma(C) = C$. By using the Proposition 23, we have $\Theta(\sigma(C)) = \overset{\circ}{v}(\Theta(C)) = \Theta(C)$. Hence $\Theta(C)$ is a set of length $8n$ over the alphabet Z_2 which is a quasi-cyclic code of index 8.

■

7 Skew cyclic DNA codes over R

In [6], the skew codes over R were studied and the Gray images of them were determined.

We will use the non trivial automorphism in [6]. For all $a + wb \in R$, it was defined by

$$\begin{aligned} \theta & : R \longrightarrow R \\ a + wb & \longmapsto a - wb \end{aligned}$$

The ring $R[x, \theta] = \{a_0 + a_1x + \dots + a_{n-1}x^{n-1} : a_i \in R, n \in N\}$ is called skew polynomial ring. It is non commutative ring. The addition in the ring $R[x, \theta]$ is the usual polynomial and multiplication is defined as $(ax^i)(bx^j) = a\theta^i(b)x^{i+j}$. The order of the automorphism θ is 2.

The following a definition and three theorems are in [6].

Definition 25 A subset C of R^n is called a skew cyclic code of length n if C satisfies the following conditions,

- i) C is a submodule of R^n ,
- ii) If $c = (c_0, c_1, \dots, c_{n-1}) \in C$, then $\sigma_\theta(c) = (\theta(c_{n-1}), \theta(c_0), \dots, \theta(c_{n-2})) \in C$

Let $f(x) + \langle x^n - 1 \rangle$ be an element in the set $\check{R}_n = R[x, \theta] / \langle x^n - 1 \rangle$ and let $r(x) \in R[x, \theta]$. Define multiplication from left as follows,

$$r(x)(f(x) + \langle x^n - 1 \rangle) = r(x)f(x) + \langle x^n - 1 \rangle$$

for any $r(x) \in R[x, \theta]$.

Theorem 26 \check{R}_n is a left $R[x, \theta]$ -module where multiplication defined as in above.

Theorem 27 A code C over R of length n is a skew cyclic code if and only if C is a left $R[x, \theta]$ -submodule of the left $R[x, \theta]$ -module \check{R}_n .

Theorem 28 Let C be a skew cyclic code over R of length n and let $f(x)$ be a polynomial in C of minimal degree. If $f(x)$ is monic polynomial, then $C = \langle f(x) \rangle$, where $f(x)$ is a right divisor of $x^n - 1$.

For all $x \in R$, we have

$$\theta(x) + \theta(\overline{x}) = 3 - 3w$$

Theorem 29 Let $C = \langle f(x) \rangle$ be a skew cyclic code over R , where $f(x)$ is a monic polynomial in C of minimal degree. If C is reversible complement, the polynomial $f(x)$ is self reciprocal and $(3 + 3w)\frac{x^n - 1}{x - 1} \in C$.

Proof. Let $C = \langle f(x) \rangle$ be a skew cyclic code over R , where $f(x)$ is a monic polynomial in C . Since $(0, 0, \dots, 0) \in C$ and C is reversible complement, we have $(\bar{0}, \bar{0}, \dots, \bar{0}) = (3 + 3w, 3 + 3w, \dots, 3 + 3w) \in C$.

Let $f(x) = 1 + a_1x + \dots + a_{t-1}x^{t-1} + x^t$. Since C is reversible complement, we have $f^{rc}(x) \in C$. That is

$$f^{rc}(x) = (3 + 3w) + (3 + 3w)x + \dots + (3 + 3w)x^{n-t-2} + (2 + 3w)x^{n-t-1} + \bar{a}_{t-1}x^{n-t} + \dots + \bar{a}_1x^{n-2} + (2 + 3w)x^{n-1}$$

Since C is a linear code, we have $f^{rc}(x) - (3 + 3w)\frac{x^n-1}{x-1} \in C$. This implies that

$$-x^{n-t-1} + (\bar{a}_{t-1} - (3 + 3w))x^{n-t} + \dots + (\bar{a}_1 - (3 + 3w))x^{n-2} - x^{n-1} \in C$$

multiplying on the right by x^{t+1-n} , we have

$$-1 + (\bar{a}_{t-1} - (3 + 3w))\theta(1)x + \dots + (\bar{a}_1 - (3 + 3w))\theta^{t-1}(1)x^{t-1} - \theta^t(1)x^t \in C$$

By using $a + \bar{a} = 3 + 3w$, we have

$$-1 - a_{t-1}x - a_{t-2}x^2 - \dots - a_1x^{t-1} - x^t = 3f^*(x) \in C$$

Since $C = \langle f(x) \rangle$, there exist $q(x) \in R[x, \theta]$ such that $3f^*(x) = q(x)f(x)$. Since $\deg f(x) = \deg f^*(x)$, we have $q(x) = 1$. Since $3f^*(x) = f(x)$, we have $f^*(x) = 3f(x)$. So, $f(x)$ is self reciprocal. ■

Theorem 30 Let $C = \langle f(x) \rangle$ be a skew cyclic code over R , where $f(x)$ is a monic polynomial in C of minimal degree. If $(3 + 3w)\frac{x^n-1}{x-1} \in C$ and $f(x)$ is self reciprocal, then C is reversible complement.

Proof. Let $f(x) = 1 + a_1x + \dots + a_{t-1}x^{t-1} + x^t$ be a monic polynomial of the minimal degree.

Let $c(x) \in C$. So, $c(x) = q(x)f(x)$, where $q(x) \in R[x, \theta]$. By using Lemma 4, we have $c^*(x) = (q(x)f(x))^* = q^*(x)f^*(x)$. Since $f(x)$ is self reciprocal, so $c^*(x) = q^*(x)e f(x)$, where $e \in Z_4 \setminus \{0\}$. Therefore $c^*(x) \in C = \langle f(x) \rangle$. Let $c(x) = c_0 + c_1x + \dots + c_tx^t \in C$. Since C is a cyclic code, we get

$$c(x)x^{n-t-1} = c_0x^{n-t-1} + c_1x^{n-t} + \dots + c_tx^{n-1} \in C$$

The vector correspond to this polynomial is

$$(0, 0, \dots, 0, c_0, c_1, \dots, c_t) \in C$$

Since $(3 + 3w, 3 + 3w, \dots, 3 + 3w) \in C$ and C linear, we have

$$\begin{aligned} (3 + 3w, 3 + 3w, \dots, 3 + 3w) - (0, 0, \dots, 0, c_0, c_1, \dots, c_t) &= (3 + 3w, \dots, 3 + 3w, \\ (3 + 3w) - c_0, \dots, (3 + 3w) - c_t) &\in C \end{aligned}$$

By using $a + \bar{a} = 3 + 3w$, we get

$$(3 + 3w, 3 + 3w, \dots, 3 + 3w, \bar{c}_0, \dots, \bar{c}_t) \in C$$

which is equal to $(c(x)^*)^{rc}$. This shows that $(c(x)^*)^{rc} = c(x)^{rc} \in C$. ■

8 DNA codes over S

Definition 31 Let f_1 and f_2 be polynomials with $\deg f_1 = t_1, \deg f_2 = t_2$ and both dividing $x^n - 1 \in R$

Let $m = \min\{n - t_1, n - t_2\}$ and $f(x) = vf_1(x) + (1 - v)f_2(x)$ over S . The set $L(f)$ is called a Γ -set where the automorphism Γ is defined as follows;

$$\begin{aligned} \Gamma &: S \longrightarrow S \\ a + wb + vc + wvd &\longmapsto a + b + w(b + d) - vc - wvd \end{aligned}$$

The set $L(f)$ is defined as $L(f) = \{E_0, E_1, \dots, E_{m-1}\}$, where

$$E_i = \begin{cases} x^i f & \text{if } i \text{ is even} \\ x^i \Gamma(f) & \text{if } i \text{ is odd} \end{cases}$$

$L(f)$ generates a linear code C over S denoted by $C = \langle f \rangle_\Gamma$. Let $f(x) = a_0 + a_1x + \dots + a_tx^t$ be over S and S -submodule generated by $L(f)$ is generated by the following matrix

$$L(f) = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_t & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \Gamma(a_0) & \Gamma(a_1) & \cdots & \cdots & \Gamma(a_t) & 0 & \cdots & \cdots & 0 \\ 0 & 0 & a_0 & a_1 & \cdots & \cdots & a_t & 0 & \cdots & 0 \\ 0 & 0 & 0 & \Gamma(a_0) & \Gamma(a_1) & \cdots & \cdots & \Gamma(a_t) & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \end{bmatrix}$$

Theorem 32 Let f_1 and f_2 be self reciprocal polynomials dividing $x^n - 1$ over R with degree t_1 and t_2 , respectively. If $f_1 = f_2$, then $f = vf_1 + (1 - v)f_2$ and $|L(f)| = 256^m$. $C = \langle L(f) \rangle$ is a linear code over S and $\Theta(C)$ is a reversible DNA code.

Proof. It is proved as in the proof of the Theorem 5 in [4]. ■

Corollary 33 Let f_1 and f_2 be self reciprocal polynomials dividing $x^n - 1$ over R and $C = \langle L(f) \rangle$ be a cyclic code over S . If $\frac{x^n - 1}{x - 1} \in C$, then $\Theta(C)$ is a reversible complement DNA code.

Example 34 Let $f_1(x) = f_2(x) = x - 1$ dividing $x^7 - 1$ over R . Hence, $C = \langle vf_1(x) + (1 - v)f_2(x) \rangle_\Gamma = \langle x - 1 \rangle_\Gamma$ is a Γ -linear code over S and $\Theta(C)$ is a reversible complement DNA code, because of $\frac{x^n - 1}{x - 1} \in C$.

9 References

- [1] Abualrub T., Ghayeb A., Zeng X., Construction of cyclic codes over $GF(4)$ for DNA computing, J. Franklin Institute, 343, 448-457, 2006.
- [2] Abualrub T., Siap I., Reversible quaternary cyclic codes, Proc. of the 9th WSEAS Int. Conference on Appl. Math., Istanbul, 441-446, 2006.

- [3] Adleman L., Molecular computation of the solution to combinatorial problems, *Science*, 266, 1021-1024, 1994.
- [4] Bayram A., Oztas E., Siap I., Codes over $F_4 + vF_4$ and some DNA applications, *Designs, Codes and Cryptography*, DOI: 10.1007/s10623-015-0100-8, 2015.
- [5] Bennenni N., Guenda K., Mesnager S., New DNA cyclic codes over rings, arXiv: 1505.06263v1, 2015.
- [6] Dertli A., Cengellenmis Y., Eren S., On the codes over the $Z_4 + wZ_4$; Self dual codes, Macwilliams identities, Cyclic, constacyclic and quasi-cyclic codes, their skew codes, *Int. J. of Foundations of Computer Science*, to be submitted, 2016.
- [7] Gaborit P., King O. D., Linear construction for DNA codes, *Theor. Computer Science*, 334, 99-113, 2005.
- [8] Guenda K., Gulliver T. A., Sole P., On cyclic DNA codes, *Proc., IEEE Int. Symp. Inform. Theory, Istanbul*, 121-125, 2013.
- [9] Guenda K., Gulliver T. A., Construction of cyclic codes over $F_2 + uF_2$ for DNA computing, *AAECC*, 24, 445-459, 2013.
- [10] Liang J., Wang L., On cyclic DNA codes over $F_2 + uF_2$, *J. Appl. Math. Comput.*, DOI: 10.1007/s12190-015-0892-8, 2015.
- [11] Ma F., Yonglin C., Jian G., On cyclic DNA codes over $F_4[u]/\langle u^2 + 1 \rangle$.
- [12] Massey J. L., Reversible codes, *Inf. Control*, 7, 369-380, 1964.
- [13] Oztas E. S., Siap I., Lifted polynomials over F_{16} and their applications to DNA codes, *Filomat*, 27, 459-466, 2013.
- [14] Pattanayak S., Singh A. K., On cyclic DNA codes over the ring $Z_4 + uZ_4$, arXiv: 1508.02015, 2015.
- [15] Pattanayak S., Singh A. K., Kumar P., DNA cyclic codes over the ring $F_2[u, v]/\langle u^2 - 1, v^3 - v, uv - vu \rangle$, arXiv:1511.03937, 2015.
- [16] Pattanayak S., Singh A. K., Construction of Cyclic DNA codes over the ring $Z_4[u]/\langle u^2 - 1 \rangle$ based on deletion distance, arXiv: 1603.04055v1, 2016.
- [17] Siap I., Abualrub T., Ghayeb A., Cyclic DNA codes over the ring $F_2[u]/(u^2 - 1)$ based on the deletion distance, *J. Franklin Institute*, 346, 731-740, 2009.
- [18] Siap I., Abualrub T., Ghayeb A., Similarity cyclic DNA codes over rings, *IEEE*, 978-1-4244-1748-3, 2008.
- [19] Yıldız B., Siap I., Cyclic DNA codes over the ring $F_2[u]/(u^4 - 1)$ and applications to DNA codes, *Comput. Math. Appl.*, 63, 1169-1176, 2012.
- [20] Zhu S., Chen X., Cyclic DNA codes over $F_2 + uF_2 + vF_2 + uvF_2$, arXiv: 1508.07113v1, 2015.